

# Computer Applications in Second Language Acquisition

Foundations for teaching, testing and  
research

*Carol A. Chapelle*

Iowa State University



**CAMBRIDGE**  
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE  
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS

The Edinburgh Building, Cambridge CB2 2RU, UK  
40 West 20th Street, New York, NY 10011-4211, USA  
477 Williamstown Road, Port Melbourne, VIC 3207, Australia  
Ruiz de Alarcón 13, 28014 Madrid, Spain  
Dock House, The Waterfront, Cape Town 8001, South Africa

<http://www.cambridge.org>

© Cambridge University Press 2001

This book is in copyright. Subject to statutory exception  
and to the provisions of relevant collective licensing agreements,  
no reproduction of any part may take place without  
the written permission of Cambridge University Press.

First published 2001

Reprinted 2001

Printed in the United Kingdom at the University Press, Cambridge

Typeset in Sabon 10.5/12pt System 3b2 [CE]

*A catalogue record for this book is available from the British Library*

*Library of Congress Cataloguing in Publication applied for*

ISBN 0 521 62637 4 hardback

ISBN 0 521 62646 3 paperback

# Contents

Series editors' preface	xiii
Thanks	xv
Abbreviations	xvii
1 Historical foundations of CASLA	1
CASLA before the microcomputer	3
The first microcomputers	8
Sophisticated microcomputers	11
<i>Computer-assisted language learning</i>	12
Software development	12
Professional issues	15
Evaluation issues	16
<i>Computer-assisted language testing</i>	17
Computer-adaptive testing	17
Alternatives to computer-adaptive testing	18
<i>Computer-assisted SLA research</i>	19
Local area networks	20
<i>Computer-assisted language learning</i>	20
<i>Computer-assisted language testing</i>	21
<i>Computer-assisted SLA research</i>	22
The Internet	23
2 The context and challenge for CASLA	27
Disciplines related to CASLA	27
<i>Educational technology</i>	28
<i>Computer-supported collaborative learning</i>	31
<i>Artificial intelligence</i>	32
<i>Computational linguistics</i>	34
<i>Corpus linguistics</i>	36
<i>Computer-assisted assessment</i>	38
From related disciplines to CASLA	40
Conclusion	43

3	Computer-assisted language learning	44
	The problem of instructed SLA	45
	<i>Cognitive conditions for SLA</i>	45
	Range of target structures	46
	The utility condition	46
	Balanced goal development	46
	Focus on form through attentional manipulation	47
	Cycles of accountability	50
	<i>Socio-affective conditions for SLA</i>	50
	<i>Other factors</i>	51
	Principles for CALL evaluation	51
	<i>Evaluation as an argument</i>	52
	<i>Judgemental and empirical analyses</i>	53
	<i>Criteria from theory and research on SLA</i>	54
	Language learning potential	55
	Learner fit	55
	Meaning focus	56
	Authenticity	56
	Positive impact	57
	Practicality	57
	<i>Criteria applied based on task purpose</i>	57
	<i>The centrality of language learning</i>	58
	Judgemental evaluation of CALL	58
	<i>Computer-assisted classroom discussion</i>	59
	<i>Microworld</i>	61
	<i>Text analysis</i>	62
	<i>Storyboard</i>	63
	<i>Concordancing</i>	64
	<i>Summary</i>	65
	Empirical evaluation of CALL	66
	<i>Evidence for CALL qualities</i>	66
	<i>Language learning potential</i>	68
	Focus on form	69
	Modified interaction	70
	Modified output	72
	Research methods	74
	<i>Learner fit</i>	80
	Level of linguistic difficulty	80
	Individual characteristics	82
	Research methods	83
	<i>Meaning focus</i>	85
	Effects of meaning-based instruction	85
	Assessing engagement with meaning	86
	<i>Authenticity</i>	88

	Comparing CALL with non-CALL activities	88
	Research methods	89
	<i>Positive impact</i>	90
	<i>Practicality</i>	93
	Conclusion	93
4	Computer-assisted language testing	95
	Test method	96
	Principles for validation in CALT	98
	<i>Validity as an argument</i>	98
	<i>Criteria for evaluating language testing</i>	100
	<i>Validation criteria and test purpose</i>	101
	<i>The centrality of construct validity</i>	103
	<i>Logical and empirical analyses</i>	103
	Test usefulness applied to CALT	105
	<i>A computer-adaptive reading test</i>	105
	<i>A computer-assisted EFL listening test</i>	108
	<i>A computer-assisted writing test</i>	109
	<i>A reading test with computer-assisted response analysis</i>	111
	<i>Summary</i>	114
	Empirical evaluation of CALT	114
	<i>Reliability</i>	116
	<i>Construct validity</i>	119
	Relationships	119
	Experimental studies	123
	Empirical task product analysis	125
	Empirical task process analysis	128
	<i>Authenticity and interactiveness</i>	129
	<i>Impact</i>	130
	<i>Practicality</i>	131
	Conclusion	131
5	Computer-assisted SLA research	132
	Investigating conditions for SLA	132
	<i>CASLR for investigating instructional conditions</i>	133
	<i>Other conditions for SLA</i>	135
	<i>Evaluation of learning condition CASLR</i>	136
	Operationalization of the learning conditions	136
	Generalizability	138
	Difficulty level	139
	Impact and practicality	140
	<i>Summary</i>	141
	Computer-assisted research tasks for assessment	141
	<i>Examples of CASLR for assessment</i>	143

	Language knowledge	143
	Strategies and processes	145
	<i>Evaluation of CASLR for assessment</i>	147
	A logical usefulness analysis	148
	Empirical usefulness analysis	149
	<i>Usefulness issues in CASLR assessments</i>	150
	Construct validity	150
	Reliability	154
	Authenticity and interactiveness	155
	Impact and practicality	155
	Conclusion	156
6	Directions for CASLA	157
	Overlapping themes	157
	<i>Reliability and learner fit</i>	158
	<i>Authenticity and generalizability</i>	159
	<i>Construct validity and operationalization of learning conditions</i>	160
	Construct validity for CALL and CASLR	160
	SLA research and CALT	161
	<i>Language learning potential and operationalization of learning conditions</i>	163
	Learning conditions for CALL	163
	CALL for operationalizing learning conditions	163
	<i>Interactiveness and meaning focus</i>	164
	<i>Positive impact</i>	165
	<i>Practicality</i>	166
	CASLA software development	166
	Needs for CASLA authoring tools	170
	<i>A means for estimating task difficulty</i>	170
	<i>Functions for analysis of learners' linguistic output</i>	170
	<i>Functions for analyzing the language of objects</i>	172
	<i>A means of organizing and creating objects</i>	172
	<i>A means for gathering process-oriented data</i>	173
	<i>A structure for learner models</i>	173
	<i>A means for authoring learning conditions</i>	174
	<i>A means for delivering CASLA via a familiar interface</i>	175
	Conclusion	175
	References	176
	Author index	205
	Subject index	211

# 1 *Historical foundations of CASLA*

At the annual TESOL convention in San Francisco in 1980, interested and curious participants attended Joan Jamieson's and my workshop introducing the use of computer software for teaching English as a second language (ESL). Joan and I had intended the workshop as a demonstration of existing ESL teaching software with an explanation of how such software is written and used in the curriculum. As newcomers to the profession, we had probably accepted uncritically the fact that the computer *was* used for teaching in the ESL program where we worked. We were therefore intrigued by questions from the audience about whether the computer *should be* used for language teaching. Various forms of this question – whether or not computers should be used for language teaching – were echoed throughout the following decade, but during the 1990s the question gradually changed from 'Should the computer be used in second language teaching?' to 'How can the computer best be used in language teaching?' As we enter the 21st century, everyday language use is so tied to technology that learning language through technology has become a fact of life with important implications for all applied linguists, particularly for those concerned with facets of second language acquisition (SLA).

Forward-looking members of the profession have suggested that the nature of communicative competence has changed in a world where communication occurs with computers and with other people through the use of computers. Writing about communicative competence in the 21st century, Rassool points out:

in a world increasingly driven by (a) the need for innovation through research and development (R&D), (b) the multilevelled changes brought about in our everyday lives as a result of the nature and speed of technological developments, (c) the volume and range of information available, and its open accessibility, (d) the multimodal features of electronic text as well as (e) its interactive nature, we require significantly *more* than just the ability to read and write in a functional way. (1999: 202; emphasis in original)

If, as Rassool suggests, 'communicative competence refers to the interactive process in which meanings are produced dynamically between information technology and the world in which we live' (Rassool, 1999: 238), language learners are entering a world in which their communicative competence will include electronic literacies, i.e., communication in registers associated with electronic communication (Murray, 2000; Warschauer, 2000).

As a consequence, anyone concerned with second language teaching and learning in the 21st century needs to grasp the nature of the unique technology-mediated tasks learners can engage in for language acquisition and how such tasks can be used for assessment. Language learners typically use computers at least to write papers, receive and send e-mail, and browse the World Wide Web; one challenge for language teachers is to shape some of their computer-using experiences into language learning experiences. To meet the challenge, the study of the features of computer-based tasks that promote learning should be a concern for teachers as well as for SLA researchers who wish to contribute to knowledge about instructed SLA. Many learners will be required to prepare for computer-assisted language tests such as those developed by the Test of English as a Foreign Language (TOEFL) program and the University of Cambridge Local Examinations Syndicate (UCLES) as well as the many Web-based language tests, including those being developed for languages of the European Union through the Diagnostic Language Assessment (DIALANG) project. Therefore, test users need to understand the issues involved in selecting such tests and helping learners prepare for them; equally critical is the knowledge of computer-assisted language testing required of test developers and researchers who construct and evaluate these new testing procedures.

To date the need for an understanding of computer-related issues in SLA has not been met by a coherent set of principles for examining past work and plotting fruitful directions. Instead, cross-disciplinary perspectives have been applied to individual efforts at development and evaluation of computer applications in second language acquisition (CASLA) – perspectives which may enrich the knowledge base concerning computer capabilities and potentials for design and evaluation. Despite the value of cross-disciplinary input, the array of computer-related methods, concepts, and initiatives presented to applied linguists can be overwhelming. Moreover, substantive progress in CASLA requires that its identity be defined, including principles for evaluation drawn from relevant work in applied linguistics. This book lays out such principles to delineate the domain of CASLA as defined through computer-assisted language learning,



computer-assisted language testing, and computer-assisted second language acquisition research. This chapter and the next begin by defining CASLA first through historical development in each of these areas and then in relation to other fields that have influenced CASLA. The following chapters focus on evaluation issues pertaining to computer applications in each area, and the final chapter suggests directions for future work on the basis of needs identified across areas.

## **CASLA before the microcomputer**

CASLA began with projects exploring development and use of computer-assisted language learning (CALL)<sup>1</sup> within the field of educational technology and was therefore shaped by perspectives in education as well as by computer hardware and software developed for purposes other than language instruction (Kerr, 1996; Saettler, 1990). In the US, computer-assisted instruction was first used in the 1950s, but examples of CALL are not documented until the 1960s, when a number of projects were undertaken to explore how the computer could be used for foreign language instruction in higher education. With a few exceptions, such projects were initiated by an individual who used computer equipment and software which had been acquired on campuses for other purposes. For example, Collett (1980), in New Zealand, reported that the idea for his French program came from a colleague in physics who had used the university's mainframe for computer-assisted instruction. Boyle, Smith, and Eckert (1976) reported a computer-based diagnostic French test also developed on a mainframe computer at a university. In the 1960s and 1970s, these small-scale individual projects, along with a few larger efforts, comprised the first experiences with CASLA.

CALL in the 1960s was supported by mainframe computers connected to terminals on a single campus or by telephone lines to terminals off campus. Computer-based learning activities, called 'courseware' were developed using programming languages and were stored on a mainframe for students to access as needed. The mainframe computers and their general-purpose programming languages of the 1970s were able to support the basic interaction

<sup>1</sup> Computer-assisted language learning (CALL) was the expression agreed upon at the 1983 TESOL convention in Toronto in a meeting of all interested participants. I have retained this term throughout this volume to refer to the area of technology and second language teaching and learning despite the fact that revisions for the term are suggested regularly.

required to implement the instructional design for this era of CALL. By today's standards, courseware was not technologically sophisticated even though it was often carefully planned. The fact that the software was stored on a single mainframe at an institution allowed for record keeping in a central location and communication among users. The mainframe also meant, however, that expenses were incurred for writing and using courseware. Because early CALL users were participating in expensive innovation, pressure existed to ensure that CALL was time well spent for learners.

Despite obstacles such as cost, individual language teachers throughout the world were fascinated by the prospects CALL appeared to offer. In the UK, for example, Rex Last and Graham Davies had each been exploring the construction of authoring software (which would simplify production of CALL) for years before they met in 1979.<sup>2</sup> Their individual experiences (e.g., Last, 1979) later became a valuable resource for an early commercial producer of language learning software in the UK. Davies' experience also made him the logical choice to head the government-funded National Centre for Computer Assisted Language Learning established in 1985.

The best-known early CALL project in North America was initiated as one part of a larger computer-assisted instruction project at Stanford University in the Institute for Mathematical Studies in the Social Sciences directed by Richard Atkinson and Patrick Suppes. The project began in collaboration with IBM, and later received funding from federal government sources. Atkinson's early research on learning foreign language vocabulary (Atkinson, 1972), still cited as having useful implications for principled design of CALL (N. C. Ellis, 1995a), was based on his mathematical learning theory rather than on then-current foreign language pedagogical practices. Atkinson (1972) found that learning, as measured by a test a week after the instruction, could be optimized significantly by having a computer program select items for practice on the basis of learners' past history of performance and item difficulty.

The work at Stanford was important also because its directors, Atkinson and Suppes, went on to form the Computer Curriculum Corporation in 1967, which continued to provide instruction in English as a second language (Saettler, 1990: 308). IBM also initiated an early project at the State University of New York at Stony Brook by funding experimental CALL materials for German (Elling, 1995).

<sup>2</sup> I am grateful to Graham Davies for the historical information he provided. For an account of past work in Europe, see Davies (1989; 1993).

Another early project began in Canada through a coordinated effort among three Ontario universities, Western Ontario, Guelph, Waterloo (and later the University of Alberta) resulting in CLEF (Computer-Assisted Learning Exercises for French), a series of 62 lessons covering basic French grammar (Paramskas, 1983), which would later be used by over 200 institutions in Canada and more abroad (Paramskas, 1995).

These are just a few of the many CALL projects that were undertaken by individuals on their university's mainframe computer during this period. Holmes and Kidd (1982) review some important ones, describing them as 'modest', emphasizing 'pedagogical principles and practical applications.' The pedagogical principles tended to go beyond the behaviorist/audio-lingual paradigms of early teaching machines by providing learners with grammatical explanations and specific feedback about their responses. For example, a German CALL project of this era at Massachusetts Institute of Technology was described as follows:

[The] tutorial to teach German reading uses the computer as a source of information to be consulted by the student as needed; the [other aspect of the program] . . . uses a model of the structure of the language being taught to enable the program to determine whether a response is correct and to provide the student with useful error analysis if it is not. (Nelson, Ward, Desch & Kaplow, 1976: 28)

The practicality and efficiency of computer use were seen as essential by instructors who were using expensive mainframe computer time. Decker (1976), for example, described his innovative approach, which involved having the computer illustrate how to perform particular grammatical operations on French learners' sentences. He then explained how this innovation would be sequenced as the first step of a process including illustration, drill, and testing to ensure that the learners had benefited. As Decker's application illustrates, and Holmes & Kidd (1982) concluded, CALL of this era was seen as a supplement to rather than as a replacement for classroom instruction. Multiple initiatives around the world explored ways in which instructional goals could be accomplished more efficiently through the use of the computer.

These projects formed the profession's initial perceptions of CALL, but what was perhaps the greatest impact on the field in this era resulted from the major commitment made in the early 1970s by the US government to support computer-assisted instruction across the curriculum. Saettler (1990) described the irony of the decision that precipitated this significant phase in the evolution of CALL.

Despite the decline of [computer-assisted instruction (CAI)], the federal government, through the National Science Foundation (NSF), decided to determine whether CAI could be made effective and available to as many teachers and schools as possible. This was the viewpoint behind the \$10 million made available in 1971 to two private companies, Control Data Corporation (CDC) and Mitre Corporation, with the idea that the two companies would compete with each other and that at least one viable CAI national system would emerge. (1990: 307)

Control Data Corporation worked with the University of Illinois at Urbana-Champaign to develop the hardware and software for the PLATO (Programmed Logic for Automatic Teaching Operations) system; the Mitre Corporation contracted with Brigham Young University in Utah to develop the TICCIT (Time-Shared, Interactive, Computer-Controlled Information Television) project. These projects, providing mainframe computer systems and software designed specifically for instruction, impacted the evolution of CALL in two ways. First, each system included major CALL components. By early 1980, TICCIT had an extensive collection of courseware that was used as an adjunct to classes in ESL, French, German, Spanish and Italian (Hendricks, Bennion & Larson, 1983) and PLATO had courseware for those languages in addition to many others such as Arabic, Chinese, Hindi, Hebrew, and Swedish (Hart, 1981a).

Second, each provided laboratories for investigation of CALL and sowed the seeds for future professional infrastructure. The TICCIT project produced a core of faculty in language teaching prepared to contribute to the evolution of technology in SLA. By the late 1970s they were pioneering videodisk technology, which resulted in one landmark project in the evolution of CALL (Schneider & Bennion, 1983). Brigham Young faculty were also leaders in computer-adaptive testing for foreign languages (e.g., Madsen, 1991). In addition, a faculty member of Brigham Young University, Frank Otto, was founder and executive director of the professional organization Computer-Assisted Language Instruction Consortium (CALICO), which has provided a forum for intellectual collaboration and growth in the field since 1984.

The PLATO project also contributed to the professional expertise in CALL. The courseware developed on that system, which supported audio (input to learners), graphics, and flexible response analysis, was the product of language teachers' best judgement of what supplemental course materials should consist of in the late 1970s. As a result of his many years of developing courseware on PLATO, Robert Hart summarized the accomplishments and identified directions for growth in 1981:

Eight years of intensive development have brought the PLATO IV grammar drill design to a high state of sophistication, so much so that further work in this direction will bring diminishing marginal returns. If we wish to make [CALL] a more powerful tool for language instruction, we really must begin to investigate qualitatively new design possibilities. (1981b: 16)

The new design possibilities he suggested were the following: (1) use of artificial intelligence techniques for analysis of learners' language in order to provide an appropriate instructional strategy, (2) diagnostic assessment of grammatical competence, (3) exploration of games and simulations which require use of 'non-trivial grammar while remaining interesting and computationally tractable' (1981b: 20), and (4) task analysis of language production, comprehension, and learning in CALL.

In retrospect, these experience-based suggestions proved to be ahead of their time. Because so few were engaged in the development and use of CALL in 1981, evolutionary progress resulting from professional discussion was not yet possible. The large majority of those who had experimented with CALL on a mainframe, or who were beginning to learn to program a microcomputer, seemed focused on the challenge of getting general-purpose hardware and software to perform for language instruction. However, primitive computer equipment and lack of professional organization were only two reasons why the early 1980s saw minimal work on these research directions. A third was perhaps that research in applied linguistics was not yet mature enough to offer principled guidance.

It would be difficult to document the many seeds sown during this period that would develop into the first attempts at computer-assisted language testing projects and computer-assisted SLA research. However, it was not an accident that early examples in the US were at Brigham Young University, where Harold Madsen and Jerry Larsen were the first in the early 1980s to report on efforts to develop computer-adaptive language testing, and the University of Illinois, where Nina Garrett began her work on computer-assisted SLA research investigating German syntax through data collected on the PLATO system (Garrett, 1982). Despite these and a few other pioneering efforts in testing and SLA research, the pre-microcomputer era of CASLA was devoted primarily to exploration of CALL.

## **The first microcomputers**

Computers became widely available to language teachers in the early 1980s. Since microcomputers did not require users to be attached to

a mainframe computer maintained by a university or business, any academic department, language school, or individual teacher could purchase one and explore its potentials for language teaching. During this period, some became interested in computer-assisted language testing (CALT) and computer-assisted second language research (CASLR), but the primary activity continued to be in CALL.

Because of the microcomputer, just three years after the inquisitive participants gathered at the San Francisco workshop, CALL had gained enough professional visibility that those working on CALL converged to discuss methodological issues, and begin formal professionalization of CALL. The 1983 annual TESOL convention in North America included papers arguing methodological issues in CALL,<sup>3</sup> and a suggestion was made to establish a professional organization (CALICO) devoted to the issues involved in language learning technology. By the following year, TESOL members were working to establish a CALL Interest Section. One year later in the UK, the British Council sponsored a course on CALL at Lancaster University which proved so popular that subsequent gatherings were organized to discuss and learn about CALL. The 1986 gathering turned out to be the founding meeting for the EuroCALL professional organization, which later received funding from the European Commission to act as a pan-European organization for CALL. In Europe, North America, and Australia, CALL's status had developed from a local curriculum or classroom issue to an international professional concern. The need was evident for teacher education through courses such as the one the British Council sponsored in 1984 at Lancaster University. In addition, a market had developed for production of introductory materials explaining computers and their classroom uses, and within a four-year period a large number of such books were published.<sup>4</sup>

By coincidence, this period overlapped the height of Steven Krashen's<sup>5</sup> popularity and hence it was fashionable to invent CALL that could be claimed to promote 'acquisition' rather than 'learning.'

<sup>3</sup> Prior to 1983, there had been only one or two sessions each year at the TESOL convention concerned with computers and language teaching.

<sup>4</sup> The following books are among those based on work of the early 1980s that were produced for teacher education: Ahmad, Corbett, Rogers, & Sussex, 1985; Brumfit, Phillips, & Skehan, 1986; Cameron, Dodd, & Rahtz, 1986; Davies, 1985; Hainline, 1987; Higgins & Johns, 1984; Hope, Taylor, & Pusack, 1984; Jones & Fortescue, 1987; Kenning & Kenning, 1983; Last, 1984; Leech & Candlin, 1986; Underwood, 1984; Wyatt, 1984.

<sup>5</sup> Krashen's view of SLA, laid out in his 1982 book, depicts two separate and unrelated processes: unconscious 'acquisition' and conscious 'learning,' the former being the most effective, in his view.

During this time, much of CALL's history was lost because what might have been the best accomplishments (e.g., perhaps Atkinson's optimal vocabulary acquisition paradigm) as well as suggested research needs (e.g., Hart's suggestion for diagnosis of grammatical competence) of the previous decades were labeled as 'learning-oriented' and therefore irrelevant to acquisition – and to CALL's future (e.g., Cook, 1985; Sanders & Kenner, 1983).<sup>6</sup> The two most influential books of this era attempted to promote CALL by explicitly attempting to dispel the idea that it must be limited to activities focusing on 'learning.' Higgins and Johns denounced the link between CALL and explicit teaching as follows:

The computer, some say, serves only the conscious process of learning, and can do nothing to facilitate acquisition . . . [W]e hope to be able to show that this view is wrong, and that the computer is quite flexible enough to serve a variety of learning theories. (1984: 17)

Underwood made the same point as follows:

It is important to stress here that this negative view [of computers as useful only for explicit learning through drills and tutorials] by no means reflects limitations in computers themselves, but rather limitations in the programs being written . . . Although much of the literature is devoted to arguing that the computer cannot do this or cannot do that, what is meant is that no one is doing it. (1984: 50)

'It' according to Underwood referred to developing 'Communicative CALL,' which he defined with 13 premises intended to be consistent with Krashen's prescriptions for creating an environment for acquisition (e.g., communicative CALL will not judge all of the language students produce). Central to Underwood's approach to creating communicative CALL was the use of techniques from artificial intelligence (i.e., natural language processing) to recognize learners' input to the computer and to generate responses in order to create a 'meaningful' conversation between computer and learner.<sup>7</sup> These two books are considered seminal works in the evolution of CALL because they supply novel ideas for CALL – programs such as games and activities based on collaborative learning – which the authors saw as providing good contexts for acquisition.

The strand of SLA research stemming from Krashen's ideas about acquisition without explicit instruction failed to provide guidance for

<sup>6</sup> At the same time, some researchers continued to work on substantive technical issues of response recognition and analysis (Pusack, 1983; Lian, 1984).

<sup>7</sup> The microcomputers widely available during the early 1980s did not have enough memory for successful implementation of the type of AI approaches (real-time written conversation) Underwood advocated.

empirically based evaluation. Evaluation of CALL tended to be comprised of the developers' or users' opinion about the extent to which an activity seemed communicative on the basis of the type of tasks it asked learners to engage in. One type of task argued to allow for communicative language practice was based on text reconstruction, which consisted of variations on cloze exercises (Higgins & Johns, 1984). Variations included the following features: words deleted on a fixed-ratio basis, words deleted on the basis of some criteria, or all words deleted;<sup>8</sup> texts that the teacher entered into the program, texts that came with the program, or texts other learners constructed; with help options and scoring, or with simple yes/no judgements concerning the correctness of the learners' entries; with the end result being the completed text, or the end result responses to comprehension questions about the text. Advocates of 'acquisition-oriented' activities saw text reconstruction as sufficiently 'communicative' and 'learner-controlled' to argue for their pedagogical value. But two factors equally instrumental in their popularity were the computational simplicity of the program required to construct such learning activities and the fact that instructors were able to input their own texts, thereby producing customized CALL materials.

Another novel invention of this era was the computer-assisted concordancer activity. Borrowed from corpus linguistics, which had already been established as a mode of inquiry in linguistics when microcomputers became widespread in the early 1980s, concordancer software is used to identify words or expressions requested by the user and display them with reference to the lines in which they occurred in a text. Higgins and Johns (1984) suggested extending the practice of concordancing to language classrooms by showing the learner how to use the concordancer to retrieve the same types of linguistic data that teachers and linguists draw from. This activity was argued to empower the learner to investigate questions of vocabulary use and grammatical collocation on their own.

Although the primary impact of SLA theory was contributed by Krashen's ideas in the early 1980s, another influence came from research on individual differences (H. D. Brown, 1980). In particular, studies looked at hypotheses from SLA about the role of individual differences on the effectiveness of different instructional approaches (Abraham, 1985) and desirability of CALL (Chapelle & Jamieson, 1986). Investigating learning style and task variables in CALL,

<sup>8</sup> Jones and Fortescue (1987) claimed that among the various text reconstruction programs, the type in which all words are deleted, called a storyboard, was the most flexible and popular.



Abraham (1985) found that field-independent learners performed better on post-tests when they had used a rule presentation (deductive) approach and field-dependent learners performed better after using a lesson presenting examples of the structure (inductive). Investigating the same learner variable, Chapelle and Jamieson (1986) found field-independent ESL students tended to have a more negative attitude toward the CALL they investigated, while the field-dependent students had more positive attitudes. Related research combined CALL with SLA through examination of learner strategies in CALL (Curtin, Avner, & Provenzano, 1981; Jamieson & Chapelle, 1987).

In short, the early 1980s was an active time in the evolution of CALL because of the diversity of ideas proposed and the growing professional discussion. This progress, however, was coupled with the regression inherent in setting aside what had come before. Loritz aptly describes the early microcomputer phase as 'the adolescence of CALL . . . a time of exploration, a time of energy and exuberance, a time when old ways are discarded, a time when new identities are born and born again' (Loritz, 1995: 47). Despite the professional visibility of the 'communicative CALL' movement, the innovative work done in the UK, and some pioneering efforts in video, many CALL developers and users during this period appeared to be reinventing the CALL of the 1970s rather than building on experience, and they did so on microcomputers which were limited in memory size and in fundamental capabilities such as audio or display of foreign language character fonts.

## **Sophisticated microcomputers**

The frustrations resulting from limitations of early microcomputers were short lived because of rapid developments during the 1980s. Throughout the decade, affordable machines came equipped with more and more memory, as well as capabilities for audio, graphics, and video. It became clear to many that computers were going to find a permanent place in language teaching and research. By the late 1980s, CALL had developed through a number of ambitious projects, professional infrastructure including teacher education, and more explicit treatment of evaluation issues. Work in CALT and CASLR began to appear as well.

## *Computer-assisted language learning*

Developments in hardware and software made tenable Hart's (1981b), Underwood's (1984) and Phillips' (1985) suggestion that

artificial intelligence in CALL be explored. Some also believed that the more sophisticated hardware and software would radically change the nature of CALL and its development:

[T]he possibilities [opened up by the more sophisticated microcomputers] are qualitatively different from those offered by the simpler equipment. It is not just a matter of having more memory to play around with: the more sophisticated machine calls for more sophisticated programming. The day of [do-it-yourself] CALL, of the hobbyist programmer, the teacher enthusiast presenting his class on Monday morning with the exercise he has spent the weekend programming, may be over. We are moving into an entirely new phase, the most distinctive feature of which is the Intelligent Tutoring System or ITS for language learning, Intelligent CALL. (Farrington, 1989: 68)

In retrospect, Farrington's characterization has been true of only one branch of CALL. The more sophisticated machine, it has turned out, can be equipped with more sophisticated software in the form of authoring tools so the 'hobbyist' and 'enthusiast' were able to communicate with the more sophisticated machines to produce more polished-looking software than had previously been possible. Even more important, however, was the fact that many CALL developers and users did not embrace the philosophy behind the intelligent tutoring system (i.e., that instruction should be designed to explicitly focus on learners' linguistic needs) and therefore pursued other CALL applications such as corpus exploration and computer-mediated communication activities.

## Software development

Others tenaciously held the goal set out by Underwood (1984) that the computer could and should be programmed to 'communicate' with the learner through the natural language processing methods developed by the computational linguists working within artificial intelligence. One instructional design using these methods was modeled after a *microworld* which is intended to create an environment for learners to explore principles of math and physics.<sup>9</sup> The software Papert (1980) designed was a computer programming language called Logo which allowed children to see geometry in action by writing commands which would instruct the computer to draw shapes and designs of their choosing. This microworld, 'math-

<sup>9</sup> In Papert's own example of a microworld for language learning, he described an activity in which children program grammatical rules into the computer, which the computer then used to create poetry. He described what is learned as an understanding of grammatical concepts.

land,' allowed children to acquire math concepts through experimentation and play in an environment which showed them the immediate effects of their mathematical statements. The idea of acquisition through manipulation of a responsive environment was attractive to CALL developers in the early 1980s who were seeking ways in which the computer could create contexts suitable for implicit 'acquisition' in Krashen's sense. Higgins and Johns (1984), for example, proposed a 'grammarland' which would 'create a miniature universe of discourse and a program which would manipulate things in that universe, answer questions about it, ask questions, or do any of these things at random if the user merely want[ed] a demonstration' (p. 75). Attempts to extend the microworld principle to CALL have taken many different forms, which vary considerably in their levels of sophistication (e.g., Ashworth & Stelovsky, 1989; Coleman, 1985; Culley, Mulford, & Milbury-Steen, 1986; DeSmedt, 1995; Durrani, 1989; Sanders & Sanders, 1995).

The software made possible by the more sophisticated microcomputers also prompted development of text analysis programs (also called grammar checkers), which were designed to provide an automatic analysis of surface features of a learner's writing and feedback about grammatical and stylistic errors. Research into text analysis had actually begun on the mainframe computers of the 1960s, when US researchers explored the capabilities of text analysis software for automatically scoring L1 English students' essays for testing (Brock, 1995; Wresch, 1993) and for providing stylistic L1 guidance to technical writers. These programs used a combination of word and phrase pattern matching and syntactic parsing to provide writers advice on how they could improve the clarity and style of their documents. In the late 1980s, similar technologies were applied to ESL learners. Dissatisfaction with the quality of the feedback provided by L1 products to L2 learners (Brock, 1993; Liou, 1991; Pennington & Brock, 1992) resulted in a number of independent efforts to develop software which would identify the types of syntactic errors that particular L2 learners make (e.g., Cook, 1988; Liou, 1991; Levin, Evans, & Gates, 1991; Loritz, 1986; Sanders, 1991).

The realities of such projects were often frustrating, as researchers attested at annual conferences, and yet the rapidly evolving technologies of this time offered high hopes for development of CALL by combining research in educational technology (particularly hypermedia), artificial intelligence, computational linguistics, and speech recognition technologies. One such vision – a CALL program for learners studying Spanish in the US – was described by Underwood (1989):

The scene is the carrel of a multi-media lab. A student is sitting in front of a color video monitor connected to stereo headphones and a tiny microphone; at her fingertips are a computer keyboard and mouse. Out of sight is a powerful computer CPU and something which looks like a CD player with a stack of 5 1/4-inch disks. Using the mouse, the student points at a little square on the screen, clicks, and the screen fills with the sights and sounds of Madrid. A voice asks her (in Spanish) if she is ready to continue; speaking into the microphone, she answers, 'Si.' She asks to talk to Javier, one of the characters she had met before, because he might have some information she needs. The screen now shows the street in front of Javier's apartment. She rings the doorbell with a mouse-click and Javier appears on the screen. 'Buenos dias,' he says. The student begins to ask him questions. At times Javier seems reluctant to talk and she must rephrase her questions to get him to respond. At other times, he says that he is sorry, but he is unable to answer such questions, for political reasons. (1989: 80)

With sights set on images such as Underwood's, a number of large CALL projects were launched during this period. The highest profile of these in the late 1980s was the industry-funded Athena Project at Massachusetts Institute of Technology in the US. The intention, as reported in 1985, was to draw upon research conducted at the Artificial Intelligence lab and other campus-wide computer resources to create a 'discovery-rich environment for the student to explore and interact with' (Kramsch, Morgenstern, & Murray, 1985: 31) through the combination of video and natural language processing technologies. In a frank retrospective account of this project, which officially ended in 1994, the emphasis on software-focused research is evident (J. H. Murray, 1995). This emphasis is underscored in a description of the natural language processing (NLP) facet of the work:

NLP is hard. When we initiated our project, we naively thought that we could successfully build an NLP system in two to three years that could analyze and respond in real time to [written] input in any one of four European languages, up to the level of a fourth semester student. Instead, it took us five years to build a system that can process second- to fourth-semester level input pretty well and often in something approaching real time . . . Grammar writing eventually expanded to fill all available time, preventing us from implementing more than prototypes of the numerous applications based on NLP that we had originally intended to create. (Felshin, 1995: 271)

The Athena project was the most ambitious, but there were a number of others of that era focused on use of either natural language processing or video. For example, researchers at the University of Delaware received funding to explore the natural language processing technologies for developing a foreign language adventure game

(Culley, Mulford, & Milbury-Steen, 1986). To produce video-based CALL materials for several foreign languages, IBM funded a consortium of universities led by the University of Iowa. Products from this project have been widely used and have served as models for subsequent work. In the UK, the Technology Enhanced Language Learning (TELL) Consortium, consisting of 15 development centers and other affiliated evaluation centers, received substantial funds for developing multimedia language learning materials. A number of small projects were also undertaken in Canada during this period (Craven, Sinyor, & Paramskas, 1990).

### Professional issues

As research and development continued in laboratories world-wide, CALL's professional infrastructure continued to expand. The Computers and Teaching Initiative Centre for Modern Languages (CTICML) was established in the UK at the University of Hull in 1988, and its journal, *ReCALL*, appeared shortly thereafter. EuroCALL continued to hold regular meetings and to seek appropriate funding – an effort which finally succeeded in 1993. Another CALL conference in Europe at the University of Exeter became a regular event and a journal based there, *Computer Assisted Language Learning: An International Journal*, appeared in 1990. In Australia the journal dedicated to CALL, *On-CALL*, appeared in the mid 1980s, and another North American journal, *CÆLL Journal*, dedicated to CALL for English as a second language, appeared in 1989.

The content of the CALL books published during this period had evolved from introductions to CALL for teachers and applied linguists who had never worked with computers to more focused treatment of a particular facet of CALL. Methodologically oriented books with practical classroom techniques continued to appear but with less introduction to the computer and more focus on the pedagogical issues of CALL (e.g., Hardisty & Windeatt, 1989; Tribble & Jones, 1990). As Farrington had predicted, some researchers probed the uses and limits of artificial intelligence in CALL (Last, 1989; Swartz & Yazdani, 1992; *Computers and the Humanities*, 1989; Bailin, 1991; Holland, Kaplan, & Sams, 1995). Others worked toward philosophies and theoretical underpinnings for CALL (Higgins, 1988; Kenning & Kenning, 1990). Perhaps the most telling indication that CALL was evolving as a professional area of concern was a more explicit treatment of evaluation issues in some publications.

## Evaluation issues

For the first time since the early CALL projects of the 1970s, explicit treatment of CALL evaluation and research issues began to appear in some CALL volumes (Dunkel, 1991; Smith, 1987; Pennington, 1989; Pennington & Stevens, 1992). Some of the suggestions that had come out as a result of experience with PLATO in the 1970s began to be taken up in a serious way. In 1981, conclusions drawn from the PLATO project were summarized as follows:

It is obvious that the developers of computer-based language materials have given far too little attention to evaluation . . . If the issues are so complex that conventional procedures (e.g., those employing group mean differences) are inappropriate for providing an answer, then we should present clear arguments why that is so and provide alternative analyses (e.g., based on individualization or optimization features). (Hart, 1981b: 16)

Several edited volumes prepared at the end of the 1980s contained papers explaining difficulties with 'conventional procedures' and laid out rationales and procedures for examining CALL from the perspectives more consistent with second language classroom research. It was clear to many at that time that 'comparative research that attempts to illustrate the superiority of computers over some other medium for language instruction should forever be abandoned' (Pederson, 1987: 125).

Alternatives to assessing technology by isolating its effects within a learning environment were drawn from work in second language classroom research (e.g., Day, 1986; Gass & Madden, 1985) and ethnographic research (e.g., Watson-Gegeo, 1988; Davis, 1995). Influenced by the qualitative classroom research tradition, CALL researchers advocated study of CALL within its larger classroom and sociocultural context. Referring to the CALL research of the prior three decades, D. Johnson (1991) noted the following:

The bulk of research on computers and learning in educational environments has focused on the cognitive aspects of learning. Yet, theory in second language acquisition and research in second language acquisition classrooms indicate that the social interactional environments of the classroom are also crucial factors that affect language learning in important ways. (1991: 62)

These suggestions from the 1980s have slowly but surely been taken up, yet the control-comparison group design seems to die hard in regular discussion at conferences and on Internet discussion lists.

### *Computer-assisted language testing*

During this period, some language testing researchers attempted to apply some of the relatively new theory and computer methods from other types of tests to language testing. Concerns were also raised that computer-assisted testing should be seen as an opportunity to extend beyond common-place types of language test items and uses.

### Computer-adaptive testing

The first-developed and most widely known use of the computer for interactive testing is a computer-adaptive test. Computer-adaptive language testing became possible through a combination of test theory for obtaining robust statistical information on test items and computer software for calculating the item statistics and providing adaptive control of item selection, presentation and evaluation (J. D. Brown, 1997; Green *et al.*, 1984; Tung, 1986; Wainer *et al.*, 1990). Harold Madsen, a professor at Brigham Young University, was among the first to apply these procedures to second language testing. He described a computer-adaptive language test as follows:

a very basic psychometric procedure which enables the examiner to measure language proficiency efficiently and with considerable precision. The adaptive or 'tailored' computer test accesses a specially calibrated item bank and is driven by a statistical routine which analyzes student responses to questions and selects items for the candidate that are of appropriate difficulty. Then, when a specified standard error of measurement is reached, the exam is terminated . . . [T]he psychometrically sound tailoring process in computer adaptive tests . . . provides for a more effective measure of language proficiency. (1991: 238–239)

The advantage was seen as primarily one of efficiency relative to paper-and-pencil tests, particularly because any individual examinee needed to complete only about one-third of the items. Moreover, Madsen reported that the international students who took the ESL tests tended to like the computer-delivered version of the test. Others seeking similar improvements and those interested in experimenting with CALT have developed similar testing projects for a variety of languages and purposes (e.g., Kaya-Carton, Carton, & Dandonoli, 1991; Burston & Monville-Burston, 1995; Brown & Iwashita, 1996; Young *et al.*, 1996; Laurier, 1999; Dunkel, 1999).

### Alternatives to computer-adaptive testing

Interest in computer-adaptive testing (CAT) was growing by the mid 1980s, but at the same time it seemed evident that this was a narrow

path to take in the exploration of CALT. In a seminal paper of this era, Canale (1986) raised questions concerning the effects of computer-adaptive reading tests on the critical needs in language testing because of the assumption of unidimensionality that their psychometric model relies on.

Such an assumption threatens to be trivializing and compromising in the following senses: First, it is overly reductionist and misleading to maintain that reading comprehension comprises only one major dimension, whatever that dimension might be . . . Second, and more generally, the assumption of unidimensionality threatens to compromise the value of CAT for educational achievement and diagnostic purposes. It is difficult to understand how CAT could serve useful achievement and diagnostic purposes if reading comprehension, for example, is assumed to be unidimensional and consequently neither influenced by instruction nor decomposable into meaningful subparts. (1986: 34–35)

Canale argued that CALT offered the opportunity to better understand multidimensional language constructs and improve the usefulness of testing for instruction, but that to realize these potentials researchers needed to look beyond testing methods constrained by a unidimensional psychometric model.

Some have argued that CALT applications could be constructed to resemble instructional activities, and that these assessments could record and analyze learners' performance to provide them with useful information about their knowledge and needs. Canale speculated on future assessments for reading comprehension by looking toward work in intelligent tutoring systems:

[W]ork on . . . 'intelligent tutoring systems' is promising for [CAT] of reading comprehension . . . Such research and images provide promising stepping stones if we are interested in moving toward more learner controlled, process-oriented and unintrusive assessment events in the language classroom. (1986: 38)

Additional suggestions about fruitful connections between instruction and assessment have been made periodically (Alderson, 1990, 1991; Corbel, 1993; Meunier, 1994; Scott & New, 1994). Alderson (1990), for example, suggested that the information gathered by the computer could encourage learners to develop their own strategies for evaluation. In fact, a number of the early CALL projects (Otto, 1989) included extensive evaluation and systematic feedback to learners. The French curriculum on the PLATO system at the University of Illinois, for example, kept records on learners' performance during each session of their work and over the course of the semester to provide them with summary information about their



performance as they requested it (Marty, 1981). However, these capabilities have yet to be explored from an assessment perspective.

### *Computer-assisted SLA research*

By the end of the 1980s, the concerns of many SLA researchers had evolved away from the idea that solely the input that learners receive through communicative activities would promote acquisition. Research through the decade had convinced many that learners need to notice and interact with linguistic input in order to acquire the target language (e.g., Schmidt & Frota, 1986; Long, 1985; Swain, 1985; Doughty & Pica, 1986). Moreover, this line of research supported empirical approaches to evaluating linguistic interaction and language development. Both of these developments in SLA research began to influence work in CASLA toward the end of the decade.

In an important paper in 1987, Doughty had laid out theoretical underpinnings from SLA with potential links to CALL. Shortly thereafter, she conducted a study using materials based on these theoretical principles, i.e., about the value of salient grammatical input. By constructing computer-assisted experimental materials that operationalized theoretically different learning conditions, Doughty (1991) compared the effects of explicitly salient L2 input with input which was not explicitly flagged to direct learners' attention. The findings, which supported theoretical predictions (e.g., learners receiving salient input performed better on grammatical post-tests than did the group receiving normal input), offered hope for the use of technology in the study of second language acquisition.

Other SLA research during this period used the computer for assessment of learners' strategies, thereby beginning to probe some of the methodological issues in assessment of processes of interest to SLA researchers such as monitoring input, advance preparation, and resourcing (Jamieson & Chapelle, 1987; Chapelle & Mizuno, 1989). These two areas – operationalization of learning conditions and assessment of learners' processes – were developed somewhat through the 1990s, but continue to hold untapped potential.

### **Local area networks**

While research and development of CASLA for microcomputers continued, the widespread use of networked computers in the early 1990s expanded the characteristics of CALL activities. By the early 1990s, many teaching staff within higher education were connected

to the Internet and had become participants in international electronic discussion lists, but the most tangible development for language learners was the adoption of Local Area Network (LAN) technology for computer labs.

### *Computer-assisted language learning*

Prior to the LAN, CALL activities had for the most part been developed around computer–learner interactions – even if more than one learner participated in those interactions at a time (e.g., Piper, 1986). LAN activities, in contrast, were built around learner–learner interactions through networked computers. Technically speaking, computer-mediated communication has been in practice since the 1960s, when users of a single mainframe computer could exchange messages in both synchronous and asynchronous modes. Only with the development of LANs and the Internet, however, was this technology put into pedagogical use for teaching collaborative L1 writing, for providing practice in second languages, and for instructing deaf learners in ‘written conversation’ (Bruce, Peyton, & Batson, 1993). Warschauer (1995b) described uses of computer-mediated communication in and across second language classrooms, and many cases are given in Warschauer (1995a).

This teaching methodology provided a written record of learners’ on-line discussion which could be examined from the perspectives of discourse analysis and SLA. Chun was among the first to publish results of such research based on an activity she constructed in which first- and second-year college learners of German in the US were to use the target language for functions associated with interpersonal communication (Chun, 1994). Through discourse analysis of the learners’ electronic discussion, she identified the variety of interpersonal functions she had hoped the activity would engender, including some she believed might not typically be found in teacher-led classrooms: initiation of discussion through questions posed by students to the rest of the class, statements to the teacher which were not in response to questions, requests for clarification, and feedback from one learner to another. The LAN-based computer-assisted discussion methodology is examined in a volume containing case-studies and discussion of research edited by Swaffar, Romano, Markley, and Arens (1998).

Other CALL research continued as well. For example, a study combining methods in educational technology and computational linguistics investigated the effects of various types of response-contingent feedback to learners of Japanese who were studying the